Remarks

A. Objection to Abstract of the Disclosure

A substitute abstract has been provided above that includes less than 150 words. No new matter has been introduced by the amendment. Thus, it is respectfully requested that this ground for objection be reconsidered and withdrawn.

B. Claim Objections

Claims 3, 7-15 were objected to as containing certain informalities such as extra spaces, erroneous punctuation and erroneous reference to steps designated by alphabetical characters.

Applicant thanks the Examiner for indicating the noted errors. The word spacing, antecedent references and punctuation of every claim has been reviewed, resulting in amendments above of virtually every claim for purposes of clarity and not to further patentably limit the scope of the invention, including the errors leading to the claim objections noted in the Action. As a result of these amendments, Applicant respectfully requests reconsideration and withdrawal of these grounds for objections.

C. Claim Rejections under 35 U.S.C. §103

Applicant respectfully submits that the Examiner has utilized hindsight in applying the cited references after reading the instant specification rather than relying upon the teachings of the cited references. Each of the rejections, that combine the cited references in various permutations, will be addressed below.

1. Claims 1, 3-4 and 15 were rejected under 35 U.S.C. §103(a) as being unpatentable over Errico[8] in view of Lyon[9].

There are several important, unobvious differences between the presently claimed invention and the processes taught by Errico and Lyon. It is clear from Figures 1 and 2, and the

[8] U.S. Patent No. 5,796,924 to Errico, *et al.*

[9] U.S. Patent No. 5,903,884 to Lyon, *et al.*

associated discussion in columns 3 and 4 of Errico that one following his method should define clusters (step 22) and only then derive classifiers for the clusters. Applicant submits that Errico's disclosure teaches away from Applicant's claimed approach (as recited in claim 1) that involves defining a classifier on the whole training set first and then using that training set to define initial clusters. In the present invention, further classifiers and clusters will be subsequently defined if needed. Additionally, Errico chooses all of his clusters at once and uses heuristics to do so. In the present invention, the training set, from which the second generation of clusters is derived experimentally using a second classifier or set of classifiers, is unknown until after the first set of clusters is defined.

These differences are nether accidental nor trivial. Rather, Applicant submits that these are two of numerous reasons why the presently claimed invention is not taught or suggested by the prior art. The methods of the presently claimed invention and Errico are fundamentally different, as the entire philosophy of cluster choice is different in the two methods. One would have to significantly alter the process of Errico to derive the present invention in at least the following unobvious ways: by replacing clusters defined *a priori* (relative to defining classifiers) to clusters defined *a posteriori*; by replacing the method step ordering (choosing clusters then deriving classifiers) with the exact opposite; by replacing the concept of defining all of the clusters at once by the concept of defining said clusters sequentially; and by replacing the heuristically defined clusters with experimentally defined clusters. These alterations of Errico surely are not an obvious extension of the teachings of Errico.

Errico does appear to teach that his classifiers may have problems that would require redefining the discriminant space, but adjusting discriminants initially selected improperly is less than an optimal approach and is avoided altogether by the presently claimed method.

Lyon is directed to a method of preventing overfitting a neural network to a finite set of training samples, by dynamically applying random distortions to the samples each time they are applied to the network during a training session. The presently claimed invention is directed to improving the generalization capability of any statistical pattern recognition system. Unlike Errico (as the Action properly notes,) Lyon allows for experimental correction of classifiers, whereas the presently claimed invention produces classifiers that require no correction or adjustment. Much of the Lyon reference is devoted to teaching ways to generate new members of the training set, whereas the presently claimed invention produces as many members as

possible before the process begins, as the currently claimed method is not subject to "over training." While it *might* be possible to follow the method teachings of Lyon to correct the clusters and/or classifiers of Errico, the presently claimed invention is not directed to and does not require such correction, expansion or updating of the clusters or classifiers.

To even further support the patentability of the claims, the following comments relating to pattern recognition are presented below.

The acknowledged leading theoretician of pattern recognition is Vladimir Vapnik[10] and his crowning accomplishment is the Support Vector Machine (SVM.) The SVM finds a separation surface with a maximum margin for a given surface complexity and a given training set. The SVM, however, does not employ an important concept of the presently claimed invention, *i.e.*, *removing correctly classified objects from the training set*, thereby setting the margin by redefining the training set. It is respectfully submitted that this is evidence that redefining the training set in accordance with the present invention is not obvious.

While other methods such as boosting, bagging and Errico's method may decompose the given training set in the search for subsets that may behave better and means to combine decisions from classifiers trained on said subsets, they persist in employing randomness or heuristics. The presently claimed invention defines the subsets experimentally on the basis of a classifier, involving no heuristics or randomness. It is respectfully submitted that the prior art's utilization of randomness or at best heuristics when an optimum choice can be found with equal ease is further evidence of unobviousness.

For at least the reasons above, Applicant respectfully submits that claims 1, 3, 4 and 15 are patentable over the cited art.

2. Claim 2 was rejected under 35 U.S.C. §103(a) as being unpatentable over Errico and Lyon in further view of Akiyama[11].

---

[10] Professor of Computer Science and Statistics at Royal Holloway University of London, and consultant to AT&T Bell Laboratories, Holmdel, NJ.

[11] U.S. Patent No. 5,602,938 to Akiyama, *et al.*

The Action points to Akiyama for support of the proposition that it is known to maximize the distance between two classes. Applicant does not dispute that at least some artisans understand that a large margin is desirable. Indeed, the SVM system noted above is precisely a way to maximize margin. However, claim 2 depends from claim 1, and Akiyama fails to make obvious the limitations of the presently claimed invention as discussed above that are neither taught nor suggested by Errico and Lyon. Thus, Applicant respectfully refers to the arguments made in section (1.) above and submits that claim 2 is similarly patentable over the cited art.

3. Claims 5-6 and 16 were rejected under 35 U.S.C. §103(a) as being unpatentable over Errico and Lyon in further view of Straforini[12], and claim 17 in further view of Akiyama.

Regarding claims 5-6 and 16-17, the arguments made in sections (1.) and (2) with respect to Errico, Lyon and Akiyama are applicable.

Straforini teaches a form of a sequential, experimental classification system directed to mixing rule-based (some might say "syntactic") with trainable pattern classifiers (some might say "statistical.") The presently claimed invention employs no rule-based classifiers.

The instant Action postulates that a pattern classifier of ordinary skill in the art would be able to determine that the training set itself can be decomposed sequentially into disjoint subsets each of which can be classified with predetermined margin with a classifier as simple as a linear discriminant. One would have to deconstruct the teachings of Errico (which asks one to consider predetermined clusters and thus teaches away from interactive cluster determination), with the knowledge that more margin means greater generalization (from Akiyama), and with Straforini (teaching that in integrating rule-based and learning-based systems one can alternate them in sequence with the latter ones dealing with items not classified by the former ones.)

Applicant respectfully suggests deriving the presently claimed invention from the cited references requires a logical leap that could only be performed in hindsight after reading the instant disclosure. One would have to abandon Errico's assumption that clusters should be assigned before classification starts. alter Straforini's teachings from alternation between modes to different generations of the same (learning-based) modes, and employ Akiyama's insight that

[12] U.S. Patent No. 6,092,059 to Straforini, *et al.*

large margin is good to determine sequentially a set of learning based rules. Abandoning the central teachings of two of the cited references and extracting other kernels that were likely not even appreciated by those inventors does not appear to support a finding of obviousness. That is, Errico did not suggest defining clusters experimentally of sequentially, Akiyama did not suggest fixing the margin as a means to define a cluster, and Straforini did not suggest improving the performance of a learning-based system by sequential subdivision into smaller disjoint sets. Applicant respectfully submits that the presently claimed invention runs counter to or require a different logical concept than the teachings of the cited references, and for this reason claims 5-6 and 16-17 are patentable over said cited references.

4. Claims 7-10 were rejected under 35 U.S.C. §103(a) as being unpatentable over Errico and Lyon in further view of Watanabe[13].

Regarding claims 7-10, the arguments made in sections (1-3) with respect to Errico and Lyon are applicable.

Watanabe teaches that multiple classifiers can be addressed in parallel by the same input data with one being selected as indicative of the class to which the input should be assigned. The presently claimed invention exhibits the potential that discriminants designed properly in sequence may (under well-defined conditions) be addressed in parallel by the input data with the class assignment being determined by the proper nonlinear logical operation on all of the outputs. There appears to be a superficial resemblance between the teachings, but selection of one output is a dramatically different proposition from the nonlinear combination of all of them. For example, the discriminant design of the presently claimed invention is dependent and sequential, while Watanabe teaches an independent and parallel design. Also, the presently claimed invention employs nonlinear combining in the decision methodology, while Watanabe teaches choosing the best response.

For the foregoing reasons and because claims 7-10 incorporate all the limitations of claim 1 by dependency, Applicant respectfully submits that the claims are similarly patentable over the cited art.

---

[13] U.S. Patent No. 5,754,681 to Watanabe, *et al.*

5. Claims 11-12 were rejected under 35 U.S.C. §103(a) as being unpatentable over Ginsburg[14] in view of Filipski[15].

Applicant does not dispute that Ginsburg teaches the fairly routine Fourier optics upon which the Action relies for support. Filipski, however, is directed to a unique and narrow range of pattern recognition problems called character recognition. In character recognition, patterns (normally 2D binary patterns such as letters or numbers) are described in terms of the presence or absence of certain defining features. Through a very complex set of operations, Filipski teaches how to determine which character is present by sequentially searching for features in an appropriate tree structure. The presently claimed invention does not employ tree structures at all, and thus there is no branching. Tree structures represent a form of sequential decision-making, whereas the presently claimed invention employs a deterministic chain without branching. What the instant disclosure and Flipski share is the insight that holistic methods are likely to fail, so sequential approaches seem indicated. But claim 11 (and 12 by dependency) recites a unique embodiment of a solution recognizing that insight that is not taught or suggested by the cited references. Thus, Applicant respectfully submits that claims 11 and 12 are patentable over said references.

6. Claims 13-14 were rejected under 35 U.S.C. §103(a) as being unpatentable over Ginsburg in view of Filipski and further in view of Watanabe.

Regarding claims 13-14, the arguments made in section (5) with respect to Ginsburg and Filipski, and section (4) with respect to the superficial relation of Watanabe to the presently claimed invention, are applicable here and are not repeated. Dependent claims 13-14 include all of the limitations of an allowable claim 11, and Applicant submits that they are similarly allowable.

Applicant respectfully submits that the above remarks clearly establish the patentability of the claimed invention over the prior art. Favorable consideration and allowance are earnestly

---

[14] U.S. Patent No. 3,993,976 to Ginsburg, Arthur P.

[15] U.S. Patent No. 4,975,975 to Filipski, Alan

solicited. Should there be any questions after reviewing this paper, the examiner is invited to contact the undersigned at 617-854-4000.

Respectfully submitted,
H. JOHN CAULFIELD,
Applicant

Dated: February 9, 2004

By: 
Jacob N. Erlich
Reg. No. 24,338
Attorney for Applicant

10050-102-RespA-020904